UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/706,546	11/12/2003	Feng-Wei Chen Russell	RSW920030186US1	6883

23550 7590 05/03/2006

HOFFMAN WARNICK & D'ALESSANDRO, LLC
75 STATE STREET
14TH FL
ALBANY, NY 12207

EXAMINER	
KOC, TARIK	

ART UNIT	PAPER NUMBER
2167	

DATE MAILED: 05/03/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/706,546	RUSSELL ET AL.	
	Examiner	Art Unit	
	Tarik C. Koc	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *12 November 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-31* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-31* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

1. Claims 1-31 are pending in this application.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1-5, 7, 16-20, 22, 24-28, 30 are rejected under 35 USC 103(a) as being obvious over . Bergamaschi et al. (Bergamaschi et al., Semantic Integration of Semistructured and Structured Data Sources, SIGMOD Record, Vol. 28, No. 1, March 1999, pp. 54-59) (hereinafter Bergamaschi), in view of Squire (Squire, Cass, Data Extraction and Transformation for the Data Warehouse, ACM, SIGMOD, 1995, pp. 446-447.),

4. Regarding claim 1, Bergamaschi discloses a computer-implemented method for mapping a schema to a disparate schema, comprising:

matching columns of a data schema to corresponding columns of another data schema to provide a mapping (page 56, column 1, paragraph 2); and

updating a matching resource based on the mapping (Section 1.1, *Generation of a Common Thesaurus,* paragraph 1; a Common Thesaurus as disclosed by Bergamaschi).

Bergamaschi does not explicitly disclose determining whether data within matching columns of the user data schema has a data type different than data within the corresponding columns of the mining model schema; and transforming the data within the matching columns of the user data schema if the data type is determined to be different; and also does not explicitly disclose wherein if the data types of matching columns are different data from the first column is transformed.

In the same field of endeavor (integration of disparate data) Squire discloses transforming data from a user data environment to a data warehouse environment (page 447, column 1, paragraph 2). Determining whether a transformation needs to be done before performing a transformation was well known in the art at the time the invention was made.

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Squire's teachings of transformation of data in a user data schema to data in a data warehouse schema with Bergamaschi's method for matching columns of a data schema to corresponding columns of another schema and updating a matching resource to obtain the claimed invention. Squire suggests in the last paragraph of column 1, page 447, that automated tools can reduce time and cost of implementation of a data warehousing system, and further suggests in the same paragraph a combination with automated tools that

"actively capture and maintain meta data related to source files... <and> transformations and mappings". Bergamaschi suggests a need in paragraph 1 of page 54 to make integration as semi-automated and scalable as possible.

Regarding claim 2, Bergamaschi and Squire do not explicitly disclose providing an opportunity to manually alter the mapping after transforming the data; and presenting a final view of the mapping after providing the opportunity, wherein the updating step is performed after the final view is presented.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik discloses providing an opportunity to manually alter the mapping after transforming the data, and presenting a final view of the mapping after providing the opportunity, wherein the updating step is performed after the view is presented (Figure 7; paragraphs 0530 and 0531; see also paragraph 0538).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of providing an opportunity to manually alter the mapping after transforming the data, presenting a final view and updating, with Bergamaschi's and Squire's suggested teachings of a computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining schema. Gorelik suggests the need for an improved automated way to discover mappings in paragraph 0002. Bergamaschi states in section 1.1, *Generation of a Common Thesaurus,* paragraph 2, that the

process of discovering mappings is a semi-automatic process wherein a designer is aided by automated tools.

Regarding claim 3, wherein the matching step comprises determining whether names of the columns of the user data schema exactly match names of the columns of the mining model data schema (Subsection *Automated extraction of relationships*, paragraph 2; see also Figure 4; Bergamaschi discloses finding synonyms, and shows in figure 4, where a synonym is inclusive of an exact match of a column name: "ID.Patient SYN CD.Patient").

Regarding claim 4, wherein the matching step further comprises determining whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource (Subsection *Automated extraction of relationships*, paragraph 2).

Regarding claim 5, wherein the matching step comprises determining whether the names of the columns of the user data schema match the names of the columns of the mining model schema based on one or more formulae (Section 1.1, *Generation of a Common Thesaurus,* paragraph 1; Bergamaschi discloses matching column names that are not exact or synonyms, but rather represent the same type of data in the disclosure of broad and narrow term relationships, which is the equivalent of formula based matching).

Regarding claim 7, wherein the matching resource is selected from the group consisting of a thesaurus, a dictionary and a similarity threshold (Section 1.1, *Generation of a Common Thesaurus,* paragraph 1).

5. Claims 6, 8, 9-15, 21, 23, 29, and 31 are rejected under 35 USC 103(a) as being obvious over . Bergamaschi et al. (Bergamaschi et al., Semantic Integration of Semistructured and Structured Data Sources, SIGMOD Record, Vol. 28, No. 1, March 1999, pp. 54-59) (hereinafter Bergamaschi), in view of Squire (Squire, Cass, Data Extraction and Transformation for the Data Warehouse, ACM, SIGMOD, 1995, pp. 446-447.), in further view of (U.S. Patent Publication #: 2005/0055369).

Regarding claim 6, Bergamaschi and Squire do not explicitly disclose wherein the matching step further comprises determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik discloses wherein the matching step further comprises determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema (paragraph 0341; see also paragraphs 0346-0349).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema, with Bergamaschi's and Squire's suggested teachings of a computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining schema. Gorelik suggests the need for an improved automated way to discover mappings in paragraph 0002. Bergamaschi suggests in paragraph 1 of page 54 a need to determine if disparate sources contain semantically related information.

Regarding claim 8, Bergamaschi and Squire do not explicitly disclose populating a schema consolidation table with names of the columns of the mining model schema, prior to the matching step; and updating the schema consolidation table with names of the matching columns of the user data schema, during the updating step.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik teaches the use of a match table (the equivalent of a schema consolidation table) indicating mappings between columns (paragraphs 0091 and 0095).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of a schema consolidation table, with Bergamaschi's and Squire's suggested teachings of a

computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining schema. Gorelik suggests the need for an improved automated way to discover mappings in paragraph 0002. Bergamaschi suggests in paragraph 2 of page 54 taking structured data sources and semistructured data into account for the purposes of capturing and reasoning about semantic aspects of schema descriptions of heterogeneous information sources for supporting integration.

6. Regarding claim 9, Bergamaschi discloses a computer-implemented method for mapping a schema to a disparate schema, comprising:

matching columns of a data schema to corresponding columns of another data schema to provide a mapping (page 56, column 1, paragraph 2); and

updating a matching resource based on the mapping (Section 1.1, *Generation of a Common Thesaurus,* paragraph 1; a Common Thesaurus as disclosed by Bergamaschi).

Bergamaschi does not explicitly disclose determining whether data within matching columns of the user data schema has a data type different than data within the corresponding columns of the mining model schema; and transforming the data within the matching columns of the user data schema if the data type is determined to be different; and also does not explicitly disclose wherein if the data types of matching columns are different data from the first column is transformed.

In the same field of endeavor (integration of disparate data) Squire discloses transforming data from a user data environment to a data warehouse environment (page 447, column 1, paragraph 2). Determining whether a transformation needs to be done before performing a transformation was well known in the art at the time the invention was made.

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Squire's teachings of transformation of data in a user data schema to data in a data warehouse schema with Bergamaschi's method for matching columns of a data schema to corresponding columns of another schema and updating a matching resource to obtain the claimed invention. Squire suggests in the last paragraph of column 1, page 447, that automated tools can reduce time and cost of implementation of a data warehousing system, and further suggests in the same paragraph a combination with automated tools that "actively capture and maintain meta data related to source files... <and> transformations and mappings". Bergamaschi suggests a need in paragraph 1 of page 54 to make integration as semi-automated and scalable as possible.

Bergamaschi and Squire do not explicitly disclose providing an opportunity to manually alter the mapping after transforming the data; and presenting a final view of the mapping after providing the opportunity, wherein the updating step is performed after the final view is presented.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik discloses providing an opportunity to manually alter the mapping after

transforming the data, and presenting a final view of the mapping after providing the opportunity, wherein the updating step is performed after the view is presented (Figure 7; paragraphs 0530 and 0531; see also paragraph 0538).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of providing an opportunity to manually alter the mapping after transforming the data, presenting a final view and updating, with Bergamaschi's and Squire's suggested teachings of a computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining schema. Gorelik suggests the need for an improved automated way to discover mappings in paragraph 0002. Bergamaschi states in section 1.1, *Generation of a Common Thesaurus,* paragraph 2, that the process of discovering mappings is a semi-automatic process wherein a designer is aided by automated tools.

Bergamaschi and Squire do not explicitly disclose populating a schema consolidation table with names of the columns of the mining model schema, prior to the matching step; and updating the schema consolidation table with names of the matching columns of the user data schema, during the updating step.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik teaches the use of a match table (the equivalent of a schema consolidation table) indicating mappings between columns (paragraphs 0091 and 0095).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of a schema consolidation table, with Bergamaschi's and Squire's suggested teachings of a computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining schema. Gorelik suggests the need for an improved automated way to discover mappings in paragraph 0002. Bergamaschi suggests in paragraph 2 of page 54 taking structured data sources and semistructured data into account for the purposes of capturing and reasoning about semantic aspects of schema descriptions of heterogeneous information sources for supporting integration.

Regarding claim 10, wherein the matching step comprises determining whether names of the columns of the user data schema exactly match names of the columns of the mining model data schema (Subsection *Automated extraction of relationships*, paragraph 2; see also Figure 4; Bergamaschi discloses finding synonyms, and shows in figure 4, where a synonym is inclusive of an exact match of a column name: "ID.Patient SYN CD.Patient").

Regarding claim 11, wherein the matching step further comprises determining whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource (Subsection *Automated extraction of relationships*, paragraph 2).

Regarding claim 12, wherein the matching step further comprises determining whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource (Subsection *Automated extraction of relationships*, paragraph 2).

Regarding claim 13, Bergamaschi and Squire do not explicitly disclose wherein the matching step further comprises determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik discloses wherein the matching step further comprises determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema (paragraph 0341; see also paragraphs 0346-0349).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema, with Bergamaschi's and Squire's suggested teachings of a computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining

schema. Gorelik suggests the need for an improved automated way to discover mappings in paragraph 0002. Bergamaschi suggests in paragraph 1 of page 54 a need to determine if disparate sources contain semantically related information.

Regarding claim 14, wherein the matching resource is selected from the group consisting of a thesaurus, a dictionary and a similarity threshold (Section 1.1, *Generation of a Common Thesaurus,* paragraph 1).

Regarding claim 15, Bergamaschi does not explicitly disclose populating a schema consolidation table with names of the columns of the mining model schema, prior to the matching step; and updating the schema consolidation table with names of the matching columns of the user data schema, during the updating step.

In the same field of endeavor (semantic discovery and mapping between data sources), Gorelik teaches the use of a match table (the equivalent of a schema consolidation table) indicating mappings between columns (paragraphs 0091 and 0095).

Accordingly, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to have incorporated Gorelik's teachings of a schema consolidation table, with Bergamaschi's and Squire's suggested teachings of a computer implemented method of mapping a user data schema to a mining model schema to obtain a system wherein an opportunity is provided to change a mapping between a user data schema and a data mining schema. Gorelik suggests the need for

an improved automated way to discover mappings in paragraph 0002. Bergamaschi suggests in paragraph 2 of page 54 taking structured data sources and semi-structured data into account for the purposes of capturing and reasoning about semantic aspects of schema descriptions of heterogeneous information sources for supporting integration.

7. Claims 16-23 are essentially the same as claims 1-8 except that they set forth the claimed invention as a system rather than a method and are rejected for the same reason as applied hereinabove.

8. Claims 24-31 are essentially the same as claims 1-8 except that they set forth the claimed invention as a product stored on a computer readable medium rather than a method and are rejected for the same reason as applied hereinabove.

Contact Information

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Tarik C. Koc whose telephone number is 571-272-6725. The examiner can normally be reached on 8:30-5:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on 571-272-4107. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Tarik C Koc

Examiner

Art Unit 2167

4/27/2006

SHAHID ALAM
PRIMARY EXAMINER